

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/785,455DATE: 02/25/97
TIME: 11:24:02

INPUT SET: S15668.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

ENTERED

SEQUENCE LISTING

(1) General Information

(i) APPLICANT

(A) NAME: Hodgson, John
(B) STREET: Great West Road
(C) CITY: Brentford
(D) STATE OR PROVINCE: Middlesex
(E) COUNTRY: United Kingdom
(F) POSTAL CODE: TW8 9EP

(A) NAME: Hodgson, John
(B) STREET: Great West Road
(C) CITY: Brentford
(D) STATE OR PROVINCE: Middlesex
(E) COUNTRY: United Kingdom
(F) POSTAL CODE: TW8 9EP

(ii) TITLE OF THE INVENTION: Novel tRNA Synthetase

(iii) NUMBER OF SEQUENCES: 2

(iv) COMPUTER-READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(v) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1974 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

RAW SEQUENCE LISTING PATENT APPLICATION US/08/785,455

DATE: 02/25/97

TIME: 11:24:07

INPUT SET: S15668.raw

```

47
48 ATGGCTAAAG AAACATTTTA TATAACAACC CCAATATACT ATCCTAGTGG GAATTTACAT 60
49 ATAGGACATG CATATTCTAC AGTGGCTGGA GATGTTATTG CAAGATATAA GAGAATGCAA 120
50 GGATATGATG TTCGTTATTT GACTGGAACG GATGAACACG GTCAAAAAAT TCAAGAAAAA 180
51 GCTCAAAAAG CTGGTAAGAC AGAAATTGAA TATTTGGATG AGATGATTGC TGGAAATTAAA 240
52 CAATTGTGGG CTAAGCTTGA AATTTCAAAT GATGATTTTA TCAGAACAAC TGAAGAACGT 300
53 CATAAACATG TCGTTGAGCA AGTGTTTGAA CGTTTATTAA AGCAAGGTGA TATCTATTTA 360
54 GGTGAATATG AAGGTTGGTA TTCTGTTCCT GATGAAACAT ACTATACAGA GTCACAATTA 420
55 GTAGACCCAC AATACGAAAA CGGTAAAATT ATTGGTGGCA AAAGTCCAGA TTCTGGACAC 480
56 GAAGTTGAAC TAGTTAAAGA AGAAAGTTAT TTCTTTAATA TTAGTAAATA TACAGACCGT 540
57 TTATTAGAGT TCTATGACCA AAATCCAGAT TTTATACAAC CACCATCAAG AAAAAATGAA 600
58 ATGATTAACA ACTTCATTAA ACCAGGACTT GCTGATTTAG CTGTTTCTCG TACATCATTT 660
59 AACTGGGGTG TCCATGTTCC GTCTAATCCA AAACATGTTG TTTATGTTTG GATTGATGCG 720
60 TTAGTTAACT ATATTTTCAGC ATTAGGCTAT TTATCAGATG ATGAGTCACT ATTTAACAAA 780
61 TACTGGCCAG CAGATATTCA TTTAATGGCT AAGGAAATTG TGCGATTCCA CTCAATTATT 840
62 TGGCCTATTT TATTGATGGC ATTAGACTTA CCGTTACCTA AAAAAGTCTT TGCACATGGT 900
63 TGGATTTTGA TGAAAGATGG AAAAATGAGT AAATCTAAAG GTAAATGTTGT AGACCCTAAT 960
64 ATTTTAAATTG ATCGCTATGG TTTAGATGCT ACACGTTATT ATCTAATGCG TGAATTACCA 1020
65 TTTGGTTCAG ATGGCGTATT TACACCTGAA GCATTTGTTG AGCGTACAAA TTTCGATCTA 1080
66 GCAAATGACT TAGGTAACCT AGTAAACCGT ACGATTTCTA TGGTTAATAA GTACTTTGAT 1140
67 GGCGAATTAC CAGCGTATCA AGGTCCACTT CATGAATTAG ATGAAGAAAT GGAAGCTATG 1200
68 GCTTTAGAAA CAGTGAAAAA CTACACTGAA AGCATGGAAA GTTTGCAATT TTCTGTGGCA 1260
69 TTATCTACGG TATGGAAGTT TATAAGTAGA ACGAATAAGT ATATTGACGA AACAACGCCCT 1320
70 TGGGTATTAG CTAAGGACGA TAGCCAAAAA GATATGTTAG GCAATGTAAT GGCTCACTTA 1380
71 GTTGAAAAATA TTCGTTATGC AGCTGTATTA TTACGTCCAT TCTTAACACA TGCGCCGAAA 1440
72 GAGATTTTTG AACAATTGAA CATAACAAT CCTCAATTTA TGGAATTTAG TAGTTTAGAG 1500
73 CAATATGGTG TGCTTACTGA GTCAATTATG GTTACTGGGC AACCTAAACC TATTTTCCCA 1560
74 AGATTGGATA GCGAAGCGGA AATTGCATAT ATCAAAGAAT CAATGCAACC GCCTGCTACT 1620
75 GAAGAGGAAA AAGAAGAGAT TCCTAGCAA CTTCAAATTG ATATTAAAGA CTTTGATAAA 1680
76 GTTGAAATTA AGGCAGCAAC GATTATTGAT GCTGAACATG TTAAGAAGTC AGATAAGCTT 1740
77 TTAAAAATTC AAGTAGACTT AGATTCTGAA CAAAGACAAA TTGTATCAGG AATTGCCAAA 1800
78 TTCTATACAC CAGATGATAT TATTGGTAAA AAAGTAGCAG TTGTTACTAA CCTGAAACCA 1860
79 GCTAAATTAA TGGGACAAAA ATCTGAAGGT ATGATATTAT CTGCTGAAAA AGATGGTGTA 1920
80 TTAACCTTAG TAAGTTTACC AAGTGCAATT CCAATGGTG CAGTGATTAA ATAA 1974

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 657 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

94 Met Ala Lys Glu Thr Phe Tyr Ile Thr Thr Pro Ile Tyr Tyr Pro Ser
95 1 5 10 15
96 Gly Asn Leu His Ile Gly His Ala Tyr Ser Thr Val Ala Gly Asp Val
97 20 25 30
98 Ile Ala Arg Tyr Lys Arg Met Gln Gly Tyr Asp Val Arg Tyr Leu Thr
99 35 40 45

```

RAW SEQUENCE LISTING PATENT APPLICATION US/08/785,455

DATE: 02/25/97
TIME: 11:24:12

INPUT SET: S15668.raw

100	Gly	Thr	Asp	Glu	His	Gly	Gln	Lys	Ile	Gln	Glu	Lys	Ala	Gln	Lys	Ala
101		50					55					60				
102	Gly	Lys	Thr	Glu	Ile	Glu	Tyr	Leu	Asp	Glu	Met	Ile	Ala	Gly	Ile	Lys
103		65					70				75					80
104	Gln	Leu	Trp	Ala	Lys	Leu	Glu	Ile	Ser	Asn	Asp	Asp	Phe	Ile	Arg	Thr
105					85						90				95	
106	Thr	Glu	Glu	Arg	His	Lys	His	Val	Val	Glu	Gln	Val	Phe	Glu	Arg	Leu
107					100						105				110	
108	Leu	Lys	Gln	Gly	Asp	Ile	Tyr	Leu	Gly	Glu	Tyr	Glu	Gly	Trp	Tyr	Ser
109			115						120				125			
110	Val	Pro	Asp	Glu	Thr	Tyr	Tyr	Thr	Glu	Ser	Gln	Leu	Val	Asp	Pro	Gln
111		130						135					140			
112	Tyr	Glu	Asn	Gly	Lys	Ile	Ile	Gly	Gly	Lys	Ser	Pro	Asp	Ser	Gly	His
113		145				150					155					160
114	Glu	Val	Glu	Leu	Val	Lys	Glu	Glu	Ser	Tyr	Phe	Phe	Asn	Ile	Ser	Lys
115					165						170				175	
116	Tyr	Thr	Asp	Arg	Leu	Leu	Glu	Phe	Tyr	Asp	Gln	Asn	Pro	Asp	Phe	Ile
117					180						185				190	
118	Gln	Pro	Pro	Ser	Arg	Lys	Asn	Glu	Met	Ile	Asn	Asn	Phe	Ile	Lys	Pro
119			195						200					205		
120	Gly	Leu	Ala	Asp	Leu	Ala	Val	Ser	Arg	Thr	Ser	Phe	Asn	Trp	Gly	Val
121		210						215					220			
122	His	Val	Pro	Ser	Asn	Pro	Lys	His	Val	Val	Tyr	Val	Trp	Ile	Asp	Ala
123		225					230					235				240
124	Leu	Val	Asn	Tyr	Ile	Ser	Ala	Leu	Gly	Tyr	Leu	Ser	Asp	Asp	Glu	Ser
125					245						250				255	
126	Leu	Phe	Asn	Lys	Tyr	Trp	Pro	Ala	Asp	Ile	His	Leu	Met	Ala	Lys	Glu
127					260						265				270	
128	Ile	Val	Arg	Phe	His	Ser	Ile	Ile	Trp	Pro	Ile	Leu	Leu	Met	Ala	Leu
129			275						280					285		
130	Asp	Leu	Pro	Leu	Pro	Lys	Lys	Val	Phe	Ala	His	Gly	Trp	Ile	Leu	Met
131		290						295					300			
132	Lys	Asp	Gly	Lys	Met	Ser	Lys	Ser	Lys	Gly	Asn	Val	Val	Asp	Pro	Asn
133		305					310					315				320
134	Ile	Leu	Ile	Asp	Arg	Tyr	Gly	Leu	Asp	Ala	Thr	Arg	Tyr	Tyr	Leu	Met
135					325						330				335	
136	Arg	Glu	Leu	Pro	Phe	Gly	Ser	Asp	Gly	Val	Phe	Thr	Pro	Glu	Ala	Phe
137					340						345				350	
138	Val	Glu	Arg	Thr	Asn	Phe	Asp	Leu	Ala	Asn	Asp	Leu	Gly	Asn	Leu	Val
139			355						360					365		
140	Asn	Arg	Thr	Ile	Ser	Met	Val	Asn	Lys	Tyr	Phe	Asp	Gly	Glu	Leu	Pro
141		370						375					380			
142	Ala	Tyr	Gln	Gly	Pro	Leu	His	Glu	Leu	Asp	Glu	Glu	Met	Glu	Ala	Met
143		385					390					395				400
144	Ala	Leu	Glu	Thr	Val	Lys	Ser	Tyr	Thr	Glu	Ser	Met	Glu	Ser	Leu	Gln
145					405						410				415	
146	Phe	Ser	Val	Ala	Leu	Ser	Thr	Val	Trp	Lys	Phe	Ile	Ser	Arg	Thr	Asn
147					420						425				430	
148	Lys	Tyr	Ile	Asp	Glu	Thr	Thr	Pro	Trp	Val	Leu	Ala	Lys	Asp	Asp	Ser
149			435						440					445		
150	Gln	Lys	Asp	Met	Leu	Gly	Asn	Val	Met	Ala	His	Leu	Val	Glu	Asn	Ile
151		450						455					460			
152	Arg	Tyr	Ala	Ala	Val	Leu	Leu	Arg	Pro	Phe	Leu	Thr	His	Ala	Pro	Lys

RAW SEQUENCE LISTING PATENT APPLICATION # US/08/785,455

DATE: 02/25/97

TIME: 14:24:18

INPUT SET: S15668.raw

153	465		470		475		480
154	Glu	Ile	Phe	Glu	Gln	Leu	Asn
155				485			490
156	Ser	Ser	Leu	Glu	Gln	Tyr	Gly
157				500			505
158	Gly	Gln	Pro	Lys	Pro	Ile	Phe
159				515			520
160	Ala	Tyr	Ile	Lys	Glu	Ser	Met
161				530			535
162	Glu	Glu	Ile	Pro	Ser	Lys	Pro
163							550
164	Val	Glu	Ile	Lys	Ala	Ala	Thr
165							565
166	Ser	Asp	Lys	Leu	Leu	Lys	Ile
167							580
168	Gln	Ile	Val	Ser	Gly	Ile	Ala
169							595
170	Gly	Lys	Lys	Val	Ala	Val	Val
171							610
172	Gly	Gln	Lys	Ser	Glu	Gly	Met
173							625
174	Leu	Thr	Leu	Val	Ser	Leu	Pro
175							645
176	Lys						
177							
178							